Instrument Review Process

Practical Information

Instrument Name

Full name of instrument and (acronym if any in parentheses)

Instrument Description

Write a description of what the instrument is typically used to measure, and if significant, what else it has been used to measure. Include constructs, number of items, and whether it was adapted from another instrument or if it is a short form version of another instrument. Reference any facts gleaned from the literature, such as differences in items from particular versions. Most of this information will be found in the developmental article.

Price

This is the price of an instrument and will be set by the publisher if it is commercially obtained. In most (not all) research situations, the author will hold rights to the instrument; therefore, contact the author for pricing or use information. If free, write "Free" (but list how in parentheses, i.e. available in the public domain, free with developer's written permission, etc.)

Administration Time

This is the time required to complete the measure; this can be found in the literature or by contacting the developer. This time may differ for different populations, versions and administration techniques (see below); therefore, provide the situation along with the time and list the appropriate reference.

Publication Year

List the four-digit year in which the instrument was developed. This is usually the publication year of the developmental article, though it may differ depending upon the author's description.

Item Readability

Some authors will provide this information in various forms, such as reading level or grade level. If this is available, then list with the appropriate reference. Otherwise, use the Flesch-Kincaid Grade Level. This is a found when using Microsoft Word's spell check (Note: This is not a default feature; it must be chosen from the Tools>Options menu under the Spelling & Grammar tab.) Type a few of the items (or copy and paste all the items, if are available electronically) in a Word document and run the spell check to obtain the Flesch-Kincaid Grade Level. Also, describe the item wording (e.g. "Most items included less than 10 words and were written in simple language" or "Most items involved two or more clauses and were written at a difficult reading level.") and mention if any items were negatively worded (e.g. "I do not feel that I am in control of my health.")

Scale Format

Explain how the items are answered: Yes/No (also known as dichotomous), Likert-type (e.g. Mostly Disagree to Mostly Agree; include number of points), ordinal (e.g. Agree to Mostly Agree), or a combination of several types of scales. Include the appropriate reference.

Administration Technique

Options include self-administered (also known as self-report) or interviewer-administered. If someone else administers it, does this person have to be a licensed clinician or can they do so with little or no training. Provide the reference, especially when test administration differed from one study to the next.

Scoring & Interpretation

Brief explanation on how the instrument is scored, whether raw scores are converted to one that is reported, any cutoffs and meanings associated with scores and score ranges. All of this information must be referenced. If there is a manual where this information can be found, then always obtain the manual. Be sure to reference it and include the manual in the folder and Annotated Bibliography.

Forms

List any other versions or language formats of the measure. This would include a version with fewer items (e.g. short form) or whether a longer form exists. Provide the reference for this information. If the reference is not used in the Annotated Bibliography, provide Pubmed-style information (i.e. author, et al., title of article, date, source.)

Annotated Bibliography

References

Use numbered references throughout the abstraction form in parentheses (e.g. (Ref: 1)) to cite the source, based on the order listed. This order is defined as 1) developmental article (or user manual, whichever was published first; 2) chronological for all remaining used articles (do not list articles that were not used in the abstraction).

The four subsections (Purpose, Sample, Methods, and Implications) are METRIC's standard format for annotating the bibliographies. Any descriptions that about samples and studies within a reference may only be described under each subsection.

Purpose: A brief statement of the problem and justification for research. Explain why the authors performed the study. For example, was the study's aim to develop a measure, validate a measure in a specific population, or was the measure used as a matter of course?

Sample: Provide information about the study sample, including descriptions of the sample demographics (e.g. sample size (N), age, gender, socioeconomic information, level of education and whether Veterans comprised all or part of the group.) If no demographic data are provided, state, "No demographic information reported.

Methods: Describe how the instrument was administered and under what conditions, as well as by whom; other study procedures that are important to the measure should be provided here, such as eligibility requirements, and gathering of information relevant to the research topic. State only the facts of the methods reported by the authors.

Implications: Do not list any psychometric data here (i.e. no statistical values). Include only information that can be generalized or conclusions the authors drew from the results. This is similar to the "Discussion" section of an article rather than the "Results" section as it pertains to the performance of this instrument. Include only limitations or implications of the instrument and not of any specific study or research topic. Again, state only the implications as the authors have reported them.

Note: Use the references presented in this ordered list for citation throughout the text using the METRIC format (Ref: #) following the sentence where data from the article has been provided.

Factors and Norms

Factor Analysis Work

If factor analysis work was performed, describe the procedure if given and whether it was confirmatory or exploratory. If it was confirmatory, how many factors did they test for? If it was exploratory, how many factors did they find?

Normative Information Availability

Provide population type with means and standard deviations.

Reliability Evidence

Test-retest

This is the correlation between at least two test administrations on the same (or similar) sample, separated by a period of time. Provide the correlation (r) values, p values, and most importantly, the time between test administrations, along with any important testing conditions. If no time interval is given, then report "No time interval given." Typically, the interval will be days or weeks, depending upon the instrument and construct. In certain instances, however, studies may report a long time interval that may bring into question whether it actually is test-retest relibaility or responsiveness (sensitivity to change). When in doubt, it is best to ask the developer or author.

Inter-rater

This is the correlation of rater's evaluation of an individual's performance that is compared to the rating of another evaluator for the same test on the same person. Kappa values may be given, and describe degree of agreement; correlation values describe the strength of the relationship; F statistics describe whether or not something is significant. If the measure is self-administered, then state this, as this reliability is not applicable (i.e. there is no rater.)

Internal Consistency

The extent to which items correlate with one another or how well they measure the same construct(s); provide correlation values and p values. An alpha coefficient is often reported for internal consistency.

Alternate Forms

This is a type of reliability in which level of agreement must be evaluated between this and other equivalent forms of a test (e.g. odd versus even items, the first half versus the second half) that measure the same construct(s). Provide r values and p values and information about the forms and testing conditions if available.

Validity Evidence

Construct/Convergent/Discriminant

These are types of validity. If there are more than one of these types discussed for a measure (especially among several articles), then group all of one type together, followed by the other, if necessary. Provide r and p values; convergent and discriminant (also called divergent) validities are tests of construct validity.

Criterion-related/Concurrent/Predictive

Also types of validity, predictive and criterion-related validity are the same and tell how well a test predicts an outside criterion; concurrent validity is in relation to a valid measure, but it must include the relation of the measure to some sort of gold standard.

Content

Another method of establishing validity based on expert judgment that the content of the measure is consistent with what is to be measured. Report r and p values.

Responsiveness Evidence

This may be given as sensitivity to change over time in the literature. The reporting of this is often confused with test-retest reliability, but the interval for the latter should be short, whereas the interval for responsiveness is much longer, depending upon the change that one sees with a given condition or treatment.

Individual-Patient Monitoring

This is also a type of validity, but it is very rarely found in the literature.

Studies where scale was used to monitor individual patients

Provide the references and any details here regarding IPM, even if no validity data was provided for the prior field.

Scale Application in VA Populations

Was this measure used in populations that included veterans. If so, include that reference in the Annotated Bibliography and list the reference number here.

Scale Application in Non-VA Populations

Was this measure used in populations that included non-veteran civilians. If so, include that reference in the Annotated Bibliography and list the reference number here.

Comments

The Comments section is reserved for the primary and secondary reviewer to express his recommendations and cautions, as well as an overview of the instrument's usefulness. METRIC has not established a specific format for this section because each reviewer has a different method of expressing opinion. Opinions are not stated anywhere in the abstraction, save for this section.

Use this guide along with METRIC Instrument Reviews, available at www.metric.research.med.va.gov to evaluate other instruments.